Self-justified equilibria: Existence and computation

Kübler, Felix ; Scheidegger, Simon

Abstract: In this paper we introduce "self-justified" equilibrium as a solution concept in stochastic general equilibrium models with a large number of heterogeneous agents. In each period agents trade in assets to maximize the sum of current utility and forecasted future utility. Current prices ensure that markets clear and agents forecast the probability distribution of future prices and consumption on the basis of current endogenous variables and the current exogenous shock. The forecasts are self-justfied in the sense that agents use forecasting functions that are optimal within a given class of functions and that can be viewed as optimally trading o˙ the accuracy of the forecast and its complexity. We show that self-justified equilibria always exist and we develop a computational method to approximate them numerically. By restricting the complexity of agents' forecasts we can solve models with a very large number of heterogeneous agents. Errors can be directly interpreted.

Self-justified equilibria: Existence and computation*

Felix Kubler
DBF, University of Zurich
and Swiss Finance Institute
fkubler@gmail.com

Simon Scheidegger
DBF, University of Zurich
simon.scheidegger@uzh.ch

February 15, 2018

Abstract

In this paper we introduce "self-justified" equilibrium as a solution concept in stochastic general equilibrium models with a large number of heterogeneous agents. In each period agents trade in assets to maximize the sum of current utility and forecasted future utility. Current prices ensure that markets clear and agents forecast the probability distribution of future prices and consumption on the basis of current endogenous variables and the current exogenous shock. The forecasts are self-justified in the sense that agents use forecasting functions that are optimal within a given class of functions and that can be viewed as optimally trading off the accuracy of the forecast and its complexity.

We show that self-justified equilibria always exist and we develop a computational method to approximate them numerically. By restricting the complexity of agents' forecasts we can solve models with a very large number of heterogeneous agents. Errors can be directly interpreted.

*Preliminary and Incomplete

1 Introduction

The assumption of rational expectations and the use of recursive methods to analyze dynamic economic models has revolutionized financial economics, macroeconomics and public finance (see e.g. Ljungqvist and Sargent (2012)). Unfortunately, for stochastic general equilibrium models with a large number heterogeneous agents rational expectations equilibria are generally not tractable, computational methods to approximate these equilibria numerically are often ad hoc, and rigorous error analysis seems impossible. Moreover, in these models, known sufficient conditions for the existence of stationary Markov equilibria are very restrictive (see Citanna and Siconolfi (2012) or Brumm et al. (2017)). In this paper we develop an alternative to rational expectations equilibria and consider temporary equilibria with forecasting functions that are optimal within a given class but that might lead to incorrect forecasts at any given time. These "self-justified"equilibria always exist and we show that by restricting the complexity of agents' forecasts one can numerically appoximate them for models with very many agents.

The basic idea of the approach is as follows. In a temporary equilibrium agents use current endogenous variables and the shock to forecast future prices and prices for commodities and assets in the current period ensure that markets clear. The forecasts are assumed to be functions that lie in a pre-specified class (a simple example are semi-algebraic functions of fixed description complexity) - the agent chooses a function to minimize a loss function of average realizations of marginal utilities along the equilibrium path and the forecasts. In the temporary equilibrium these expectations might be far from correct and agents might make large mistakes. However, they forecasts are optimal given the agents' constraints. The concept does not require identical expectations or identical forecasts across agents. Different types of agents can have different expectations and different forecasting functions.

To prove existence of self-justified equilibrium we make the simplifying assumption that accounting is finite. That is to say, we assume that beginning-of-period portfolios across agents lie on some finite (arbitrarily fine) grid and that agents' portfolio-choices in the current period induce a probability distribution over this grid. This assumption can be viewed as a technical approximation to a continuous model, but one can also think of bounded rationality justifications. For example, one might to assume that at the beginning of a period an agent cannot measure his financial wealth with arbitrary precision and makes small errors in rounding. In any case, while the assumption is necessary for the technical argument is has no effect on the computed solutions since all computations are necessarily using finite precision arithmetics.

There is a large and diverse body of work exploring deviations from rational expectation (see, e.g., Spear (1989), Sargent (1993), Guesnerie (2001,2005), Gabaix (2014), Adam et al. (2016)). Much of this work is motivated by insights from behavioral economics about agents' behavior or by the search for simple economic mechansism that enrich the observable implications of standard

models. The motivation of this paper is rather different in that we want to develop a simple alternative to rational expectations that allows researchers to rigorously analyze stochastic dynamic models with heterogeneous agents.

As Sargent (1993) points out, "when implemented numerically ... rational expectations models impute more knowledge to the agent within the model ... than is possessed by an econometrician", and a sensible approach to relax rational expectations is "expelling rational agents from our model environment and replacing them with 'artificially intelligent 'agents how behave like econometricians."This quote emodies the idea underlying self-justified equilibria – in order to construct a tractable model of the macro-economy that takes into account substantial heterogeneity across agents one needs to assume that agents' expectations can actually be computed by the modeler.

Applied dynamic general equilibrium modeling has been critized for its failure to take into account the large heterogeneity in tastes and technologies across agents. Farmer and Foley (2009) make this point forefully and strongly advocate the use of so called agent-based models to understand macro-economy dynamics. An agent-based model is a computerized simulation of a number of decision-makers and institutions, which interact through prescribed rules. The agents can be as diverse as needed but, as Baptista et al. (2016) point out, in these agent based models "behavioral rules are often arbitrary". Up to now it seemed to complicated to incorporate substantial heterogeneity into in large-scale dynamic GE models because existing solution methods are not able to handle this amount of heterogeneity. Using the concept of self-justified equilbria, one can incorporate large scale heterogenity into general equilibrium models, potentially improve their usefulness for applied work and bridge the gap between agent based modeling and applied general equilibrium.

The rest of the paper is organized as follows. In Section 2 the general economy is introduced.

2 A general dynamic Markovian economy

We consider a Bewley-style overlapping generations model (see Bewley (1984)) with incomplete financial markets and a continuum of agents. Time is indexed by $t \in \mathbb{N}_0$. Exogenous shocks z_t realize in a complete, separable metric space $\mathbf{Z}$, and follow a first-order Markov process with transition probability $\mathbb{P}(.|z)$ defined on the Borel σ-algebra $\mathcal{Z}$ on $\mathbf{Z}$—that is, $\mathbb{P} : \mathbf{Z} \times \mathcal{Z} \to [0,1]$. Let $(z_t)_{t=0}^{\infty}$, or in short (z_t), denote this stochastic process and let $(\mathcal{F}_t)$ denote its natural filtration (i.e., the smallest filtration such that (z_t) is $\mathcal{F}_t$-adapted). A history of shocks up to some date t is denoted by $z^t = (z_0, z_1, \ldots, z_t)$ and called a date event. Whenever convenient, we simply use t instead of z^t.

At each date event a continuum of ex ante identical agents enter the economy, live for A periods, and differ ex post by the realization of their idiosyncratic shocks. Each agent faces idiosyncratic shocks, $y_1, ..., y_A$, that have support in a finite set $\mathbf{Y}^A$. We denote by $\eta_{y^a}(y_{a+1})$ the (conditional) probability of idiosyncratic shock y_{a+1} for an agent with shock history y^a, $\eta_0(y_1)$ to denote the

probability of idiosyncratic shock y_1 at the beginning of life, and, $\eta(y^a)$ to denote the probability of a history of idiosyncratic shocks. We assume that the idiosyncratic shocks are independent of the aggregate shock, that they are identically distributed across agents within each type and age and, as in the construction in Proposition 2 in Feldman and Gilles (1985), that they "cancel out" in the aggregate, that is, the joint distribution of idiosyncratic shocks within a type ensures that at each history of aggregate shocks, z^t, for any $y^a \in \mathbf{Y}^a$ the fraction of agents with history $y^a = (y_1, ..., y_a)$ is $\eta(y^a)$. This allows the focus on equilibria for which prices and aggregate quantities only depend on the history of aggregate shocks, z^t. I denote the set of all date events at time t by $\mathbf{Z}^t$ and, taking z_0 as fixed, I write $z^t \in \mathbf{Z}^t$ for any $t \in \mathbb{N}_0$ (including $t = 0$). At each z^t there are finitely many different agents actively trading (distinguishing themselves by age and history of shocks), who are collected in a set $\mathbf{I} = \cup_{a=1}^{A} \mathbf{Y}^a$. A specific agent at a given node z^t is denoted by $y^a \in \mathbf{I}$.

At each date event there is a single perishable commodity, the individual endowments are denoted by $e_{y^a}(z^t) \in \mathbb{R}_+$ and assumed to be time-invariant and measurable functions of the current aggregate shock[1]. Each agent who can be identified by his date-event of birth, z^t, has a time-separable expected utility function

$$U_{z^t}((x_{t+a})_{a=0}^{A-1}) = \mathbb{E}_t \left[\sum_{a=1}^{A} u_{y^a} \left(x_{z^t, t+a-1} \right) \right],$$

where $x_{z^t,t+a-1} \in \mathbb{R}_+$ denotes the agent's (stochastic) consumption at date $t + a - 1$.

There are J assets, $j \in \mathbf{J} = \{1, \ldots, J\}$ traded at each date event. Assets can be infinitely lived Lucas trees in unit net supply or one-period financial assets in zero net supply. The net supply of an asset j is denoted by $\bar{\theta}_j \in \{0, 1\}$. Assets are traded at prices q and their (non-negative) payoffs depend on the aggregate shock and possibly on the current prices of the assets $f_j : \mathbb{R}_+^J \times \mathbf{Z} \to \mathbb{R}_+$. If asset j is a Lucas tree (i.e., an asset in positive net supply), then $f_j(q, z) = q_j + d_j(z)$ for some dividends $d_j : \mathbf{Z} \to \mathbb{R}_+$. Asset j could also be a collateralized loan whose payoff depends on the value of the underlying collateral, or an option, or simply a risk-free asset. The aggregate dividends of the trees are defined as $d(z_t) = \bar{\theta} \cdot f(q(z^t), z_t) - \bar{\theta} \cdot q(z^t)$. An agent y^a faces trading constraints $\theta \in \mathbf{\Theta}_{y^a} \subset \mathbb{R}^J$, where $\mathbf{\Theta}_{y^A} = \{0\}$ for all $y^A \in \mathbf{Y}^A$. To simplify notation we write $\vec{\theta} = (\theta_{y^a})_{y^a \in \mathbf{I}}$, $\vec{\theta^-} = (\theta^-_{y^a})_{y^a \in \mathbf{I}}$ and $\vec{x} = (x_{y^a})_{y^a \in \mathbf{I}}$.

It is useful to define the set of possible portfolio holdings with market-clearing built-in as

$$\mathbf{\Theta} = \{\vec{\theta} : \sum_{y^a \in \mathbf{I}} \eta(y^a)\theta_{y^a} = \bar{\theta}, \quad \theta_{y^a} \in \mathbf{\Theta}_{y^a} \text{ for all } y^{a-1} \in \mathbf{I}\}.$$

Similarly let the set of all beginning-of-period portfolio holdings be

$$\mathbf{\Theta}^- = \{\vec{\theta^-} : \theta^-_{y^1} = 0, \sum_{y^{a-1} \in \mathbf{I}} \eta(y^{a-1})\theta^-_{y^a} = \bar{\theta} \text{ and } \theta^-_{y^a} \in \mathbf{\Theta}_{y^{a-1}} \text{ for all } y^a\}.$$

[1] As opposed to the standard formulation where an agent's fundamentals are functions of his current idiosyncratic shock, y, we assume that they are functions of the history of all shocks - clearly these formulations are equivalent if one allows for a sufficiently rich set $\mathbf{Y}$.

2.1 Finite accounting

The first crucial and non-standard assumption of the paper is that accounting is finite, i.e. actual beginning of period portfolios lie in a finite set. This assumption will allow us to obtain simple existence results below but it comes at the cost of some opaqueness.

We assume that the space of exogenous shocks can be decomposed into two complete, separable metric spaces, $\mathbf{Z} = \mathbf{Z}_0 \times \mathbf{Z}_1$ with Borel σ-algebra $\mathcal{Z} = \mathcal{Z}_0 \otimes \mathcal{Z}_1$, and the shock is given by $z = (z_0, z_1)$. We assume that z_0 is independent across time and independent of z_1, and that z_0 does not affect fundamentals other than a rounding shock ϵ which is only determined by z_0.

We let $\widehat{\mathbf{\Theta}}^- \subset \mathbf{\Theta}^-$ be a finite set and assume that given $\vec{\theta}(z^t)$, we have

$$\vec{\theta}^-(z^{t+1}) \in \arg \min_{\vec{\theta}^- \in \widehat{\mathbf{\Theta}}^-} \|\bar{\theta} + \epsilon(z_{t+1}) - \theta^-\|_2,$$

with $\bar{\theta}_{y^a} = \theta_{y^{a-1}}$ for all $a = 2, \dots A$, $y^a \in \mathbf{Y}^a$ and $\bar{\theta}_{y^1} = 0$ for all $y^1 \in \mathbf{Y}$. In this formulation $\epsilon(z_0)$ should be interpreted as a (small) rounding error.

Since we assume that ϵ_0 is iid and since it only affects the current rounding error we can define the aggregate state space to only contain the z_1 shock. We then denote by $\mathbf{S} = \mathbf{Z}_1 \times \widehat{\mathbf{\Theta}}^-$ the (aggregate) state-space.

Note that can think of the rounding error alternatively as follows. A profile of choices in the current period, $\vec{\theta} \in \mathbf{\Theta}$, given current shock $z \in \mathbf{Z}$, induces a probability distribution $\mu_{z,\vec{\theta}}$ over states in the subsequent period $s \in \mathbf{S}$. We denote by G the number of elements in $\widehat{\mathbf{\Theta}}^-$ and by $\mathbf{\Delta}^{ZG-1}$ the $ZG-1$ dimensional unit simplex in $\mathbb{R}^{ZG}$. This allows us to write $\mu_{z,\vec{\theta}} \in \mathbf{\Delta}^{ZG-1}$ and, in a slight abuse of notation $\mu : \mathbf{Z} \times \mathbf{\Theta} \to \mathbf{\Delta}^{ZG-1}$.

Assuming finite accounting has several justifications. As stated we assume that agents take the fact that beginning-of-period portfolios always lie on a finite grid as a technological constraint. Alternatively we could assume that actual portfolios lie in $\mathbf{\Theta}^-$ but that agent cannot measure their wealth arbirarily finely and make their decisions based on rounded values, exhibiting some degree of bounded rationality. This would lead to similar results, but it turns out to be simpler to view the finiteness of $\mathbf{\Theta}^-$ as a trading constraint. From a practical point of view the assumption seems innocuous. Because of finite precision arithmetic in scientific computations and numerical method will lead to $\vec{\theta}^-$ lying on a (possibly very fine) grid. However, from a technical point, the assumption turns out to be crucial. It is not clear which of our results hold true in the limit as the grid becomes dense in $\mathbf{\Theta}^-$. Throughout it is assumed that the support of $\epsilon(.)$ is centered around zero and very small.

Unfortuantely, finite accounting makes the maximization problem of the agent very complicated as the optimal solution cannot be characterized by first order conditions. In fact in order to define the maximization problem formally we would need to specify the conditional probability of and individual's portfolio on the grid, given the choices of all other agents. Instead we assume that the

agent solves a continuous and convex problem. Given the transition of the aggregate state $\mathbb{Q}(s'|s)$ each agent y^a's value function is a function of the aggregate state and his individual portolfio's and it is given by

$$\begin{aligned} V_{y^a}(s,\theta^-) &= \max_{\theta\in\mathbf{\Theta}_{y^a},x\in\mathbb{R}_+} u_{y^a}(x) + \sum_{s'\in\mathbf{S}} \mathbb{Q}(s'|s) \sum_{y_{a+1}\in\mathbf{Y}} \eta_{y^a}(y_{a+1}) V_{y^a}(s',\theta_{y^a}) \\ &\text{s.t. } x + \theta\cdot q(s) = f(q(s),z)\cdot\theta^- + e_{y^a}(z) \end{aligned} \tag{1}$$

The agent is bounded rational as the solution to this problem will not be identical to the actual optimal solution. However, if $\widetilde{\mathbf{\Theta}}$ is sufficiently fine the solutions will be very close.

2.2 Recursive equilibria

It is useful to start our analysis by defining a recursive equilibrium with rational expectations. In this context almost rational expectations means that agents correctly forecast all future prices conditional on shocks, but, and therefore the "almost", they make (arbitrarily small) systematic mistakes in that they think that the portfolio they choose in the current period carries over to the subsequent period (i.e. while they understand that states are finite they do not understand that their own portfolio gets randomly rounded to lie on the the finite grid).

The formal definition is as follows.

DEFINITION 1 *A recursive equilibrium consists of a function $\rho_q : \mathbf{S} \to \mathbb{R}_+^J$, $\rho_{\vec{\theta}} : \mathbf{S} \to \mathbb{R}^J$ and a transition $\mathbb{Q}(s'|s)$ and well as value functions for all agent $y^a \in \mathbf{I}$ that depend on the aggregate state s, and the agents' beginning of period portfolio $\theta^-_{y^a}$, i.e. $V_{y^a} : \mathbf{S}\times\mathbf{\Theta}_{y,a} \to \mathbb{R}$, such that*

- *Each agent y^a solves his maximization problem, for all $s \in \mathbf{S}$*

$$\begin{aligned} \rho_{\theta_{y^a}}(s) &= \arg\max_{\theta\in\mathbf{\Theta}_{y^a},x\in\mathbb{R}_+} u_{y^a}(x) + \sum_{s'\in\mathbf{S}} \mathbb{Q}(s'|s) \sum_{y_{a+1}\in\mathbf{Y}} \eta_{y^a}(y_{a+1}) V_{y^{a+1}}(s',\theta_{y^a}) \\ &\text{s.t. } x + \theta\cdot q(s) = f(q(s),z)\cdot\theta^-_{y^a} + e_{y^a}(z), \end{aligned}$$

  where $V_{y^a}(.)$ solve (1).

- *Markets clear, i.e.*

$$\sum_{y^a} \eta(y^a)\rho_{\theta_{y^a}}(s) = \bar{\theta}, \text{ for all } s \in \mathbf{S}$$

- *Agents forecast future states correctly*

$$\mathbb{Q}(s'|s) = \mu_{z,\rho_{\vec{\theta}(s)}}(s')$$

3 Self justified equilibria

In a competitive environment agents choose asset-holdings in the current period to maximize expected utility and current prices ensure that markets clear. In order to understand how today's asset choices affect future utilities the agent needs to form some expectations about future prices and compute his optimal life-cycle asset-holdings under these prices. In the definition of a recursive equilbirium it is assumed that the agent knows the state transition and the map from states to prices. For economies with a very large number of agents this seems unrealistic. It certainly leads to a situation where equilibrium cannot be approximated numerically. In order to relax this assumption it turns out to be useful to model the forecasting of prices and the recursive solution of the agents' problem in one step and assume that the agents' expectations are over the next period's marginal utility of asset holdings. We first reformulate a recursive equilibrium in this framework and we then allow these forecasts to be imperfect and heterogneous across agents.

Each agent, $y^a \in \mathbf{I}$, is characterized by a functions $M_{y^a} : \mathbf{Z} \times \mathbf{\Theta} \to \mathbb{R}^J_+$ that predicts marginal utilities of assets in the next period on the basis of the current portfolio-holdings across agents and the current shock. We denote by $\mathbf{M} = \times_{y^a \in \mathbf{I}} \mathbf{M}_{y^a}$ and write $\vec{M} = (M_{y^a})_{y^a \in \mathbf{I}}$. Throughout we assume that $M_{y^A}(.) = 0$ for all $y^A \in \mathbf{Y}^A$, forecasts of agents of age A are irrelevant.

Assuming concavity of utility, the first order conditions are necessary and sufficient for agents' optimality and we can write an agent y^a's maximization problem as

$$\begin{aligned} \max_{x \in \mathbb{R}_+, \theta \in \Theta_{y^a}} \; & u_{y^a}(x) + M_{y^a}(z, \vec{\theta}_{y^a}) \cdot \theta \text{ s.t.} \\ & x + \theta \cdot \bar{q} - e_{y^a}(z) - \theta^-_{y^a} \cdot f(\bar{q}, z) \geqslant 0. \end{aligned}$$

The agent takes as given current average portfolio choices across all agents, $\vec{\theta}$ and current prices q. The function $M_{y^a}(.)$ is part of the agent's characteristics - we impose rationality assumptions on this, depending on the equilibrium concept.

We define the temporary equilibrium correspondence

$$\mathbf{N} : \mathbf{S} \times \mathbf{M} \rightrightarrows \mathbb{R}^I_+ \times \mathbf{\Theta} \times \mathbb{R}^J$$

as

$$\begin{aligned} \mathbf{N}(s, \vec{M}) \quad = \quad & \{(\bar{x}_{y^a}, \bar{\theta}_{y^a})_{y^a \in \mathbf{I}}, \bar{q}) \in \mathbb{R}^I_+ \times \mathbf{\Theta} \times \mathbb{R}^J_+ : \\ & (\bar{x}_{y^a}, \bar{\theta}_{y^a}) \in \arg \max_{x_{y^a} \in \mathbb{R}_+, \theta_{y^a} \in \Theta_{y^a}} u_{y^a}(x_{y^a}) + M_{y^a}(z, (\bar{\theta}_{y^a})_{y^a \in \mathbf{I}}) \cdot \theta_{y^a} \text{ s.t.} \\ & x_{y^a} + \theta_{y^a} \cdot \bar{q} - e_{y^a}(z) - \theta^-_{y^a} \cdot f(\bar{q}, z) \geqslant 0 \quad \text{for all } y^a \in \mathbf{I}\}. \end{aligned} \tag{2}$$

Assuming that for a given $\vec{M}$ the set $\mathbf{N}(s, \vec{M})$ is non-empty for all $s \in \mathbf{S}$ there exists a single valued selection $N(s; \vec{M})$ because $\mathbf{S}$ is finite. We write

$$N(s; \vec{M}) = \left(N_{\vec{x}}(s; \vec{M}), N_{\vec{\theta}}(s; \vec{M}), N_q(s; \vec{M}) \right).$$

Before defining a self-justified equilibrium it useful to derive the functions $\vec{M}(.)$ for the case of a recursive (rational expectations) equilibrium. Given a recursive equilibrium with value functions $V_{y^a}(.)$ and policy functions ρ there must a selection $N(.,\vec{M})$ of $\mathbf{N}(.,\vec{M})$ such that for all $s \in \mathbf{S}$, if

$$M_{y,a}(z, N_{\vec{\theta}}(s;\vec{M})) = \sum_{s' \in \mathbf{S}} \mu_{z,N_{\vec{\theta}}(s;\vec{M})}(s') \sum_{y^{a+1} \in \mathbf{Y}} \eta_{y^a}(y^{a+1}) \frac{\partial V_{y^{a+1}}}{\partial \theta}(\theta_{y^a}, s)$$

The crucial innovation of this paper is to allow for hetergeneous and possibly incorrect forecasts across agents. We assume that the agents deviate from rational expectations with respect to one crucial aspects: They cannot evaluate (or store) arbitrary functions but instead approximate the equilibrium forecasts by "simple " functions. These functions could be simple because they aggregate $\vec{\theta}$ into a lower dimensional vector, or because they belong to some convenient class of functions - a simple example would be semi-algebraic functions of fixed description complexity, we give another example in the computational section below.

Each agents y^a considers a set of functions $\mathbf{M}_{y^a}$ to choose from, has an infinite amount of information on equlibrium allocations and prices and minimizes the weighted sum of a loss-function $L_{y^a}(.)$ to obtain his forcasting function M_{y^a}. We assume that the loss function is a function of the realized marginal utility, the forecasted marginal utility and the forecasting function, $L_{y_a} : \mathbb{R}^J + \times \mathbb{R}^J_+ \times \mathbf{M}_{y^a} \to \mathbb{R}$ to allow for the possibility that the agent trades off between more accurate forecasts and more complex forecasting functions. The formal definition of a self-justified equilibirum is as follows.

DEFINITION 2 *A self-justified equilibrium consists of forecasts $\vec{M} \in \mathbf{M}$, a selection $N(.;\vec{M})$ of $\mathbf{N}(.,\vec{M})$, and a measure $\mathbb{P}$ on $\mathbf{S}$, such that*

1. *$\mathbb{P}$ is invariant given the law of motion induced by $N(.)$, that is to say for all $s' \in \mathbf{S}$*

$$\mathbb{P}(s') = \sum_{s \in \mathbf{S}} \mathbb{P}(s) \mu_{N_{\vec{\theta}}(s;\vec{M})}(s')$$

2. *For each y^a, $a < A$, M_{y^a} provides the best average approximation under this measure, i.e.*

$$M_{y^a} \in \arg\min_{M \in \mathbf{M}_{y^a}} \sum_{s \in \mathbf{S}} \mathbb{P}(s) L_{y^a}\left(M(z, N_{\vec{\theta}}(s;\vec{M})), m(s), M(.)\right),$$

where

$$m(s) = \sum_{s' \in \mathbf{S}} \mu_{z,N_{\vec{\theta}}(s;\vec{M})}(s') f(N_q(s';\vec{M})) \sum_{y^{a+1} \in \mathbf{Y}} \eta_{y^a}(y^{a+1}) u'_{y^{a+1}}(N_{x_{y^{a+1}}}(s';\vec{M}))$$

Similarly to the concept of "self-confirming"equilbirium (see e.g. Fudenberg and Levine (1993) or Cho and Sargent (2008)) a self-justified equilibrium can be interpreted as the outcome of a learning-process which itself is not modeled in the theory. The crucial difference is that in a self-justified equilibrium an agent's forecasts can be incorrect in every step, as long as they are the best forecasts the agent can choose.

Note that in this definition the value function does not occur and $\frac{\partial V_{y^{a+1}}}{\partial \theta}(\theta_{y^a}, s)$ is replaced by $u'_{y^{a+1}}(N_{x_{y^{a+1}}}(s'; \vec{M}))$. This is motivated by the idea that a self-justified equilibrium is a stationary point of a learning process where agents learn to forecasts prices as well as their own value function. Since consumption is only observed on the finite set $\mathbf{S}$ we assume that agents approximate the derivative of their value function at θ_{y^a} by the derivative of the value function at $\theta^-_{y^{a+1}}$. If the grid $\widehat{\mathbf{\Theta}}^-$ is sufficiently fine this difference is like to be tiny. By the envelope theorem

$$\frac{\partial V_{y^{a+1}}}{\partial \theta}(\theta^-_{y^{a+1}}, s) = u'_{y^{a+1}}(N_{x_{y^{a+1}}}(s'; \vec{M})).$$

4 Existence

Througout the paper we make the following assumptions on fundamentals.

Assumption 1

1. *For each $y^a \in \mathbf{I}$ the Bernoulli-utility function $u_{y^a}(.)$ is continuously differentiable, strictly increasing, strictly concave, and satisfies an Inada conditions*

   $$u'_{y^a}(x) \to \infty \text{ as } x \to 0,$$

   individual Endowments are positive, i.e.,

   $$e_{y^a}(z) > 0 \text{ for all } z \in \mathbf{Z}.$$

2. *The set $\mathbf{\Theta}$ is compact and for each $y^a \in \mathbf{I}$ the set $\mathbf{\Theta}_{y^a}$ is a closed convex cones containing $\mathbb{R}^J_+$.*

3. *The payoff functions, $f : \mathbb{R}^J_+ \times \mathbf{Z} \to \mathbb{R}^J$ are non-negative valued and continuous. Moreover, for any $i, j = 1, \ldots, J$ the payoff $f_j(q)$ only depends on q_i if $\bar{\theta}_i > 0$.*

4. *For all $\theta^- \in \Theta_{y^{a-1}}$*

   $$\theta^-_{y^a} \cdot f(q, z) \geqslant 0 \text{ for all } q \in \mathbb{R}^J_+, z \in \mathbf{Z}.$$

5. *The function $\mu : \mathbf{Z} \times \mathbf{\Theta} \to \mathbf{\Delta}^{ZG-1}$ is continuous in $\vec{\theta}$ for all $z \in \mathbf{Z}$.*

Assumptions 1.1-1.3 are standard. Assumption 1.4 is motivated by collateral and default. Constraints ensure that agents cannot borrow against future endowments. In our formulation this is true independently of prices – we implicitly allow for default (see e.g. Kubler and Schmedders (2003)). Assumption 1.5 is satisfied if $\epsilon(.)$ has a continuous probability density function.

Since we assumed Θ^- to be finite, for fixed $\vec{M} \in \mathbf{M}$ a selection of the temporary equilibrium correspondence can be viewed as a vector $N \in \left(\mathbb{R}^I_+ \times \mathbf{\Theta} \times \mathbb{R}^J_+\right)^{GZ}$. We make the following reduced-form assumption on forecasting- and loss functions.

Assumption 2

1. *For all* $\mu \in \Delta^{ZG-1}$ *and all* $N \in \left(\mathbb{R}_+^I \times \mathbf{\Theta} \times \mathbb{R}_+^J\right)^{GZ}$ *the following*

$$\widetilde{M}_{y^a}(N,\mu) = \arg\min_{M \in \mathbf{M}_{y^a}} \sum_{s \in \mathbf{S}} \mu(s) L_{y^a}\left(M(z, N_\theta(s)), m(s), M(.)\right),$$

$$m(s) = \sum_{s' \in \mathbf{S}} \mu_{N_\theta(s)}(s') f(N_q(s')) \sum_{y^{a+1} \in \mathbf{Y}} \eta_{y^a}(y^{a+1}) u'_{y^{a+1}}(N_{x_{y^{a+1}}}(s')),$$

is well defined (i.e. the arg min *exists and is unique). The function* $\widetilde{M}_{y^a}(N,\mu,(z,\vec{\theta}_{y^a}))$ *is jointly continuous in* $(N,\mu,\vec{\theta})$ *for all* $z \in \mathbf{Z}$.

2. *All functions in* $\mathbf{M}$ *are uniformly bounded above, i.e. there is some* $\bar{m}$ *such that*

$$M_{y^a,j}(z,\vec{\theta}) < \bar{m} \text{ for all } z \in \mathbf{Z}, \vec{\theta} \in \mathbf{\Theta}, j \in \mathbf{J} \text{ and } y^a \in \mathbf{I}$$

Assumption 2.1 is quite strong since it rules out arbitrarily optimistic expecations. Without further specifying $\mathbf{M}$ and $L(.)$ it is needed for some of our results, however.

With these assumptions existence of a self-justified equilibrium simply reduces to the existence of a finite-dimensional fixed point, in particular - the main result of this section is as follows.

THEOREM 1 *Under Assumptions 1-2 there exists a self-justified equilibrium.*

Proof. We decompose the economy into sub-economies for each $s \in \mathbf{S}$. Assumption 1.3 implies that there exist l, r such that whenever $\vec{\theta} \in \mathbf{\Theta}$,

$$l < \theta_{y^a,j} < r \text{ for all } y^a \in \mathbf{I}, j \in \mathbf{J}.$$

Let $\mathbf{T} = [l,r]^J$ and let $\mathbf{X} = [0, \max_{z \in \mathbf{Z} y^a \in \mathbf{I}} \frac{e(z)+d(z)}{\eta(y^a)}]$. Assumption 2.2 guarantees that any forecast is bounded above by $\bar{m}$. Let $\mathbf{O} = [0,\bar{m}]^{JGZ}$.

We construct a upper-hemi-continuous, non-empty and convex-valued correspondence, $\mathbf{\Phi}^{\bar{q}}$, mapping $(\mathbf{X}^I \times \mathbf{T}^I \times \Delta^J)^{GZ} \times \mathbf{\Delta}^{GZ} \times \mathbf{O}^I$ to itself which has a fixed point. We then argue that this implies the existence of a self-justified equilibrium.

For all $y^a \in \mathbf{I}$ and all $s \in \mathbf{S}$ let

$$\begin{aligned} \mathbf{\Phi}_{y^a,s}(M_s,(p_s,q_s)) \quad = \quad & \arg\max_{x \in \mathbf{X}, \theta \in \mathbf{\Theta}_{y^a} \cap \mathbf{T}} u_{y^a}(x) + M_s \cdot \theta \text{ s.t.} \\ & (x_{y^a} - e_{y^a}(z)) + \theta_{y^a} \cdot \frac{1}{p_s} q_s - \theta^-_{y^a} \cdot f(\frac{1}{p_s} q_s, z) \leqslant 0 \end{aligned}$$

and let

$$\mathbf{\Phi}_{0,s}(\vec{\theta}_s, \vec{x}_s) = \arg\max_{(p,q) \in \mathbf{\Delta}^J} p(\sum_{y^a \in \mathbf{I}} \eta(y^a)(x_{y^a,s} - e_{y^a}(z) - d(z))) + q \cdot (\sum_{y^a \in \mathbf{I}} \eta(y^a)(\theta_{y^a,s} - \bar{\theta}))$$

Let

$$\mathbf{\Phi}_\mu((\vec{\theta}_s)_{s \in \mathbf{S}}) = (\mu(s) \sum_{s' \in \mathbf{S}} \mu_{z,\vec{\theta}_s}(s'))_{s \in \mathbf{S}}$$

and let

$$\mathbf{\Phi}_{M_{y^a}}(\mu,(\vec{x}_s,\vec{\theta}_s,p_s,q_s)_{s\in\mathbf{S}})=(\widetilde{M}_{y^a}(s))_{s\in\mathbf{S}},$$

where

$$\widetilde{M}_{y^a}=\arg\min_{M\in\mathbf{M}_{y^a}}\sum_{s\in\mathbf{S}}\mu(s)L_{y^a}\left(M(z,\vec{\theta}_s),m(s),M(.)\right),$$

with

$$m(s)=\sum_{s'\in\mathbf{S}}\mu_{z,\vec{\theta}_s}(s')f(\frac{1}{p_{s'}}q_{s'})\sum_{y^{a+1}\in\mathbf{Y}}\eta_{y^a}(y^{a+1})u'_{y^{a+1}}(x_{y^{a+1}}(s')).$$

Assumptions 1 and 2 guarantee that the mapping $\mathbf{\Phi}=\times_{s\in\mathbf{S}}\left((\times_{y^a\in\mathbf{I}}\mathbf{\Phi}_{y^a,s})\times\mathbf{\Phi}_{0,s}\right)\times\mathbf{\Phi}_{M_{y^a}}\times\mathbf{\Phi}_{\mu}$,

$$\mathbf{\Phi}:(\mathbf{X}\times\mathbf{T}\times\mathbf{\Delta}^J)^{GZ}\times\mathbf{\Delta}^{GZ-1}\times\mathbf{O}^I\rightrightarrows(\mathbf{X}\times\mathbf{T}\times\mathbf{\Delta}^J)^{GZ}\times\mathbf{\Delta}^{GZ-1}\times\mathbf{O}^I$$

is non-empty and convex valued and upper hemi-continuous. By Kakutani's fixed point theorem there exists a fixed point. Assumption 1 guarantees that the additional constraints imposed by $\mathbf{T}$ and $\mathbf{X}$ are not binding, and hence this fixed point is a self-justified equilbirium. □

Using the exact same proof strategy we can prove the existence of a self-justified equilibrium where forecasts are correct.

PROPOSITION 1 *There is a $\vec{M}$ and a selection $N(.)$ of $\mathbf{N}$ such that for each y^a, M_{y^a} provides a correct forecast, i.e. for all $s\in\mathbf{S}$*

$$M(z,N_{\vec{\theta}}(s;\vec{M}))=m(s)$$

where

$$m(s)=\sum_{s'\in\mathbf{S}}\mu_{z,N_{\vec{\theta}}(s;\vec{M})}(s')f(N_q(s';\vec{M}))\sum_{y^{a+1}\in\mathbf{Y}}\eta_{y^a}(y^{a+1})u'_{y^{a+1}}(N_{x_{y^{a+1}}}(s';\vec{M}))$$

The proof is analogous to the proof of Theorem 1. The only difference is that we do not need to proof the existence of an invariant distribution and $\mathbf{\Phi}_{M_{y^a}}(\mu,(\vec{x}_s,\vec{\theta}_s,p_s,q_s)_{s\in\mathbf{S}})$ is simply set equal to the expected marginal utility, i.e.

The discretization of the state-space enables us to prove a very simple result – without this, strong assumptions are needed to ensure the existence of a recursive rational expectations equilibrium (see Brumm et al. (2017)).

5 A simple introductory examples

In order to illustrate the concept of self-justified equilibria and to describe our general computational strategy below it is useful to focus on a specific simple example. In the simplest example we assume that agents live for 60 periods and that there is a single agent per generation. Agents can trade in a single Lucas-tree and in Arrow securities. In our framework it is useful to assume that the Arrow-securities pay in the Lucas-tree (as in Gottardi and Kubler (2015)). In this example we make

the strong assumption that agents only take their own asset holdings to forecast future M and that their forecasts a linear.

It turns out that for simple calibrations of the model forecasts in a self-justified equilibrium are extremely accurate.

Assuming that agents form simple linear forecasts using only their own asset holdings, an agent y^a's chooses $\theta_{y^a} \in \mathbb{R}_+^Z$ to solve

$$\max_{\theta \in \mathbb{R}_+^S, x \geqslant 0} u_{y^a}(x) + \sum_{z' \in \mathbf{Z}} (a_{y^a,z,z'} + b_{y^a,z,z'} \bar{\theta}_z$$

$$\text{s.t. } e_{y^a}(z) + \theta^-_{y^a,z}(q + d(z)) - \sum_{z'} \theta_{y^a,z'} q(z,z')$$

Where $\bar{\theta}_z$ equals the chosen $\theta_{y^a,z}$

Clearly we make two strong assumptions in this example. First we assume that agents only use their own choices to forecasts future marginal utilities. This implies that prices are just a function of the shock and as it would be the case, choices depend on asset positions.

Second the agents have linear forecasts. This is obviously incorrect for agents of age $A-1$, their consumption is linear in asset holdings, by concavity their marginal utility is non-linear. However, it is not clear how much asset holdings vary in a self-justified equilibrium. As it turns out a linear approximation is sometimes accurate.

5.1 The computational strategy

In this simple setup the computation of self-justified equilibria is very simple and reduces to linear regression and the repeated solution of non-linear systme of equations. In particular we employ an iterative scheme to solve for the optimal forecasting functions.

It is useful to define the temporary equilibrium system of inequalities as the system of all agents' KKT-conditions together with the Z market clearing conditions, i.e.

$$F((\theta_{y^a}), q, \kappa_{y^a} | \theta^-_{y^a}) = \begin{cases} -u'_{y^a}(e_{y^a}(z) + \theta^-(q + d(z) - q \cdot \theta) q_z + \beta M_{y^a}(z, z', \theta_{y^a}) + \kappa_{y^a,z'} & \text{for all } y^a, z' \\ \kappa_{y^a} \cdot \theta_{y^a} & \\ \sum_{y^a} \theta_{y^a}(z') = 0, & \text{for all } z' \in \mathbf{Z} \end{cases} \tag{3}$$

The details of the algorithm are then as follows:

1. Make an initial guess for each agent's forecasting

$$M^0_{y^a,z'} : \mathbf{Z} \times \mathbb{R}_+ \to \mathbb{R}_+$$

with

$$M^0_{y^a,z'}(z, \theta_z) = a_{y^a,z,z'} + b_{y^a,z,z'} \theta_z.$$

Choose an approximation accuracy $\bar{\eta}$.

2. Make one time iteration step:

   (a) For $NSIM$ periods, simulate a noisy temporary equilibrium for givne forecasts M^0. That is to say starting with some z, θ^- solve the system (3). using pseudo random numbers draw a new z' and set $\theta^-_{y^a} = \theta_{y^{a-1},z'} + \epsilon$.

   (b) Regress the equilibrium values of $(q_z + d(z))u'(x)$ on θ_z to obtain new coefficients $a_{y^a,z,z'}, b_{y^a,z,z'}$ and a new forecastig function M^1

3. If

$$\|M^1 - M^0\| < \eta$$

   then set $M^* = M^1$ and terminate the algorithm. Else set $M^0 = M^1$ and repeat time iteration step.

The temporary equilibrium needs to be noisy for the computations to work. Since many young agents never choose to save in equilibrium we cannot determine their forecasts for regions where they would save. Therefore, with low probability, we have that agents wake up with assets although they never saved.

5.2 A simple self-justified equilibrium with accurate forecasts

For the simplest example, assume that $A = 60$, $Z = 2$ and that there is a single agent per generation. The agent has CRRA utility function with $u_{y^a}(c) = \beta^a \frac{c^{1-\gamma}}{1-\gamma}$. We take $\beta = 0.96$, $\gamma = 2$. Individual endowments are

$$e_a(1) = 0.4 + a/500, e_a(2) = 0.9 * (0.4 + a/500) \text{ for } a < 50$$

$$e_a(1) = e_a(2) = 0.4 \text{ for } a \geqslant 50.$$

Also assume that $d(z) = 2$ for both $z = 1, 2$. and that $\pi(1) = \pi(2) = \frac{1}{2}$.

In the computed self-justified equilibrium forecasting errors are tiny, they are the smallest for young agents (around 0.01 percent) and the largest for old agents (around 0.5 percent). Average errors across agents are about 0.1 percent.

This result is of course consistent with many examples in the literature where one finds pseudo aggregation (most notably Krusell and Smith (1996)) and Chien et al. (2011)). The main reason why the simple forecasts fare well in this example is that there is almost no variation in asset holdings. In Figure 1 we show the forecasts of an agent of age 5 in shock 1 for next period's shock 1. As one can see in the FIgure, there is only a tiny variation in asset holdings and realized marginal utiliies and forecasted marginal utiliies more or less coincide.

[FIGURE 1 ABOUT HERE]

One motivation for self-justified equilibrium is to find a way to systematically make agents more rational if needed. For this we first need to construct a simple example where forecasts that do not take into account the wealth distribution across agents do not do a very good job.

5.3 Moving away from the simple example

As it turns out, the result that linear forecasts are quite accurate holds true for a wide variety of parameter specifications. One simple case where this breaks down can be obtained by assuming that agents across generations have different subjective beliefs over the aggregate shocks. While this does not really fit our model and does not fit the idea that the agents know invariant distributions it gives us a simple model to compare different algorithms.

Concretely we modify the simple example above in assuming that agents of ages 1-10 and of ages 50-59 have incorrect probabilities in that

$$\pi^a(1) = 0.6, \pi^a(2) = 0.4 \text{ for } a = 1, \dots 10$$

$$\pi^a(1) = 0.4, \pi^a(2) = 0.6 \text{ for } a = 50, \dots 59.$$

All other agents have the correct beliefs.

With this specification forecasts a systematically misspecified not only because they are linear but mainly because future marginal utilities for asset holding do not only depend on own choices. Figure 2 depicts the same forecasting function as Figure 1, but for this specification with heterogeneous beliefs. Clearly linear functions do not do a good job, but moreover, it is clear that other variables have to be added to make forecasts good.

[Figure 2 about here]

In the next section we give an overview over the approximation method we use. We then show how our general computational strategy is a straightforward generalization of the simple method described in this section.

6 Function approximation on high-dimensional and irregularly-shaped domains

To solve for self-justified equlibria in general, we need to repeatedly approximate and interpolate multi-variate policy function on irregularly-shaped—that is, non-hypercubic domains. In such environments, standard grid-based methods such as "Smolyak" (see, e.g., Krueger and Kubler (2004) and Judd et al. (2014)) or "adaptive sparse grids" (see, e.g., Brumm and Scheidegger (2017) and Brumm et al. (2015)), will fail. To this end, we will follow closely Scheidegger and Bilionis (2017) and use Gaussian process regression (GPR) (see, e.g., Rasmussen and Williams (2005) and Sec. 6.1)

in combination with active subspaces (see, e.g., Constantine et al. (2014) and Sec. 6.2) and Bayesian Gaussian mixture models (see, e.g., Rasmussen (2000) and Sec. 6.3).

6.1 Gaussian process regression

Below, we provide a very brief introduction to Gaussian process regression (GPR) based on Rasmussen and Williams (2005) and Scheidegger and Bilionis (2017), and references therein.

GPR is a nonparametric regression method from supervised machine learning, and addresses the problem of learning input–output mappings from observed data—the so-called training set. In our case, observations stem from a computer code. Given a data set $\mathcal{D} = \{\left(\mathbf{x}^{(i)}, t^{(i)}\right) | i = 1, ..., \mu\}$ consisting of μ input vectors $\mathbf{x}^{(i)} \in \mathbb{R}^N$ and corresponding, potentially noisy, observations $t^{(i)} = f(\mathbf{x}^{(i)}) + \epsilon, \epsilon \sim \mathcal{N}\left(0, \sigma_\epsilon^2\right)$, we want to deduce a model of the unknown function f that generated the data such that we then can make predictions for new inputs $\mathbf{x}^*$ that we have not seen in the training set. The matrix

$$\mathbf{X} = \left\{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(\mu)}\right\} \tag{4}$$

is commonly referred to as *training inputs*, and

$$\mathbf{t} = \left\{t^{(1)}, \ldots, t^{(\mu)}\right\} \tag{5}$$

is the vector of the corresponding *training targets*—that is, observations.

To enable a prediction—that is, interpolation based on information contained in $\mathcal{D}$, we must make assumptions about the characteristics of the underlying functions, as GPR is a Bayesian regression method. We start by defining a probability measure on the function space, where $f(\cdot)$ lives corresponding to our beliefs. Before seeing any data, we model our state of knowledge about $f(\cdot)$ by assigning a GP prior to it. We say that $f(\cdot)$ is a GP with *mean function* $m(\cdot; \phi)$ and *covariance function* $k(\cdot, \cdot; \phi)$, and write

$$f(\cdot)|\phi \sim \mathrm{GP}(f(\cdot)|m(\cdot; \phi), k(\cdot, \cdot; \phi)), \tag{6}$$

where $\phi \in \mathbf{\Theta} \subset \mathbb{R}^{d_\theta}$ and $d_\theta \in \mathbb{N}$ are the so-called *hyper-parameters* of the model. The covariance matrix can be chosen, but must be positive semi-definite and symmetric. Througout our work, we use the *square exponential* (SE)

$$k_{\mathrm{SE}}(\mathbf{x}, \mathbf{x}'; \phi) = s^2 \exp\left\{-\frac{1}{2}\sum_{i=1}^{N}\frac{(x_i - x_i')^2}{\ell_i^2}\right\}, \tag{7}$$

where $\phi = \{s, \ell_1, \ldots, \ell_N\}$, with $s > 0$ being the variability of the latent function V, and $\ell_i > 0$ the characteristic lengthscale of the i-th input. The hyper-parameters of the covariance function are typically estimated by maximizing the likelihood (see Scheidegger and Bilionis (2017), and references

therein). Given a set of training inputs $\mathbf{X}$, Eq. 6 induces a Gaussian prior on the corresponding response outputs:

$$\mathbf{f} = \left\{ f\left(\mathbf{x}^{(1)}\right), \ldots, f\left(\mathbf{x}^{(\mu)}\right) \right\}. \tag{8}$$

In particular, $\mathbf{V}$ is distributed as

$$\mathbf{f}|\mathbf{X}, \phi \sim \mathcal{N}\left(|\mathbf{m}, \mathbf{K}\right), \tag{9}$$

where $\mathcal{N}\left(\cdot|\mathbf{m}, \mathbf{K}\right)$ is the PDF of a multivariate Gaussian random variable with $\mathbf{m} := \mathbf{m}(\mathbf{X}; \phi) \in \mathbb{R}^{\mu}$ being the mean function evaluated at all points in $\mathbf{X}$, and $\mathbf{K} \in \mathbb{R}^{\mu \times \mu}$ is the covariance matrix with $K_{ij} = k(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}; \phi)$ (see Eq. 7). In the Bayesian framework we are operating in, we have to model explicitly the measurement process that gives rise to the observations $\mathbf{t}$. We do so by assuming that measurements are independent of one another, and that they are normally distributed about $f(\cdot)$ with variance s_n^2:

$$t^{(i)}|f\left(\mathbf{x}^{(i)}\right), s_n \sim \mathcal{N}\left(t^{(i)}\middle| f\left(\mathbf{x}^{(i)}\right), s_n^2\right), \tag{10}$$

where $s_n > 0$ is an additional hyper-parameter that must be determined from the training targets. Using the independence of the observations, we get

$$\mathbf{t}|\mathbf{f}, s_n \sim \mathcal{N}\left(\mathbf{t}\middle|\mathbf{f}, s_n^2 \mathbf{I}_N\right). \tag{11}$$

The *likelihood* of the observations is

$$\mathbf{t}|\mathbf{X}, \phi, s_n \sim \mathcal{N}\left(\mathbf{t}\middle|\mathbf{m}, \mathbf{K} + s_n^2 \mathbf{I}_N\right). \tag{12}$$

Bayes's rule combines the prior GP (see Eq. 6) with the likelihood (see Eq. 12) and yields the *posterior* GP

$$f(\cdot)|\mathbf{X}, \mathbf{t}, \phi, s_n \sim \mathrm{GP}\left(f(\cdot)\middle|\tilde{m}(\cdot), \tilde{k}(\cdot, \cdot)\right), \tag{13}$$

where the *posterior* mean and covariance functions are given by

$$\tilde{m}(\mathbf{x}) := \tilde{m}(\mathbf{x}; \phi) = m(\mathbf{x}; \phi) + \mathbf{K}(\mathbf{x}, \mathbf{X}; \phi)\left(\mathbf{K} + s_n^2 \mathbf{I}_N\right)^{-1}(\mathbf{t} - \mathbf{m}) \tag{14}$$

and

$$\begin{aligned} \tilde{k}(\mathbf{x}, \mathbf{x}') &:= \tilde{k}(\mathbf{x}, \mathbf{x}'; \phi, s_n) \\ &= k(\mathbf{x}, \mathbf{x}'; \phi) - \mathbf{K}(\mathbf{x}, \mathbf{X}; \phi)\left(\mathbf{K} + s_n^2 \mathbf{I}_N\right)^{-1} \mathbf{K}(\mathbf{X}, \mathbf{x}; \phi), \end{aligned} \tag{15}$$

respectively. To carry out interpolation tasks when iterating on policies, one has to work with the predictive (marginal) distribution of the function value $f(\mathbf{x}^*)$ for a single test input $\mathbf{x}^*$ conditional on the hyper-parameters ϕ and s_n—namely,

$$f(\mathbf{x}^*)|\mathbf{X}, \mathbf{t}, \phi, s_n \sim \mathcal{N}\left(f(\mathbf{x}^*)|\tilde{m}(\mathbf{x}^*), \tilde{\sigma}(\mathbf{x}^*)\right), \tag{16}$$

where $\tilde{m}(\mathbf{x}^*) = \tilde{m}(\mathbf{x}^*; \phi)$ is the *predictive mean* given by Eq. 14, and $\tilde{\sigma}^2(\mathbf{x}^*) := \tilde{k}(\mathbf{x}^*, \mathbf{x}^*; \phi, s_n)$ is the *predictive variance*. Note that the predictive mean can be used as interpolation value.

6.2 Active subspaces

Standard GPs are not able to deal with very high input dimensions. This is because they rely on the Euclidean distance to define input-space correlations. Since the Euclidean distance becomes uninformative as the dimensionality of the input space increases (Bengio et al., 2005), the number of observations required to learn the function grows enormously. To this end, following Scheidegger and Bilionis (2017), we couple GPs to active subspaces (see, e.g., Constantine et al. (2014)).

One way of dealing with this problem is to discover and exploit structures that reduce the dimensionality of the input space. Specifically, we assume that the response surface can be well approximated with the following form:

$$f(\mathbf{x}) \approx h\left(\mathbf{W}^T\mathbf{x}\right), \tag{17}$$

where the matrix $\mathbf{W} \in \mathbb{R}^{D\times d}$ projects the high-dimensional input space, $\mathbb{R}^D$, to the low-dimensional *active subspace*, $\mathbb{R}^d, d \ll D$, and $h : \mathbb{R}^d \to \mathbb{R}$ is a d-dimensional function known as the *link* function. Note that the representation of Eq. 17 is not unique. All matrices $\mathbf{W}$ whose columns span the same subspace of $\mathbb{R}^d$ yield identical approximations. Thus, without loss of generality, we restrict our attention to matrices with orthogonal columns. An added benefit of enforcing this orthogonality is that the columns of $\mathbf{W}$ correspond to directions of the input space on which the response is most sensitive. Mathematically, we write $\mathbf{W} \in V_d\left(\mathbb{R}^D\right)$, where $V_d\left(\mathbb{R}^D\right)$ is the set of $D \times d$ matrices with orthogonal columns,

$$V_d\left(\mathbb{R}^D\right) := \left\{\mathbf{A} \in \mathbb{R}^{D\times d} : \mathbf{A}^T\mathbf{A} = \mathbf{I}_d\right\}, \tag{18}$$

with $\mathbf{I}_d$ the $d \times d$ unit matrix. $V_d\left(\mathbb{R}^D\right)$ is also known as the *Stiefel manifold*. If d is much smaller than D, then the problem of learning the surrogate is significantly simplified.

The standard way to find the active subspace requires using gradient information (Constantine, 2015; Lukaczyk et al., 2014; Dow and Wang, 2013), that is, in addition to training data, we need the respective gradients

$$\mathbf{G} = \left\{\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(N)}\right\}, \tag{19}$$

where

$$\mathbf{g}^{(i)} = \nabla f\left(\mathbf{x}^{(i)}\right) \in \mathbb{R}^D \tag{20}$$

and $\nabla f(\cdot)$ is the gradient of $f(\cdot)$,

$$\nabla f(\cdot) = \left(\frac{\partial f(\cdot)}{\partial x_1}, \ldots, \frac{\partial f(\cdot)}{\partial x_D}\right). \tag{21}$$

The approach here operates in two steps. First, it identifies the projection matrix $\mathbf{W} \in V_d\left(\mathbb{R}^D\right)$ using gradient information. Second, it projects all inputs to the active subspace and then applies GP regression to learn the map between the projected inputs and the output.

Let $\rho(\mathbf{x})$ be a PDF on the input space such as the PDF of a uniform random variable, and define the matrix

$$\mathbf{C} := \int (\nabla f(\mathbf{x}))(\nabla f(\mathbf{x}))^T \rho(\mathbf{x}) d\mathbf{x}. \tag{22}$$

Since $\mathbf{C}$ is symmetric positive definite, it admits the form

$$\mathbf{C} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T, \tag{23}$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \cdots, \lambda_D)$ is a diagonal matrix containing the eigenvalues of $\mathbf{C}$ in decreasing order, $\lambda_1 \geqslant \cdots \geqslant \lambda_D \geqslant 0$, and $\mathbf{V} \in \mathbb{R}^{D \times D}$ is an orthonormal matrix whose columns correspond to the eigenvectors of $\mathbf{C}$. The classical AS approach suggests separating the d largest eigenvalues from the rest,

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2 \end{bmatrix},$$

(here $\mathbf{\Lambda}_1 = \text{diag}(\lambda_1, \ldots, \lambda_d), \mathbf{V}_1 = [\mathbf{v}_{11} \ldots \mathbf{v}_{1d}]$, and $\mathbf{\Lambda}_2, \mathbf{V}_2$ are defined analogously), and setting the projection matrix to

$$\mathbf{W} = \mathbf{V}_1. \tag{24}$$

Intuitively, $\mathbf{V}$ rotates the input space so that the directions associated with the largest eigenvalues correspond to directions of maximal function variability (Constantine, 2015).

It is impossible to evaluate Eq. 22 exactly. Instead, the usual practice is to approximate the integral via MC, that is, assuming that the observed inputs are drawn from $\rho(\mathbf{x})$, one approximates $\mathbf{C}$ using the observed gradients by

$$\mathbf{C}_N = \frac{1}{N} \sum_{i=1}^{N} \mathbf{g}^{(i)} \left(\mathbf{g}^{(i)}\right)^T. \tag{25}$$

In practice, the eigenvalues and eigenvectors of $\mathbf{C}_N$ are found using the singular value decomposition (Golub and Van Loan, 1996) of $\mathbf{C}_N$. The dimensionality d is determined by looking for sharp drops in the spectrum of $\mathbf{C}_N$ (Constantine, 2015), or by looking at the Bayesian information criterion.

6.3 Bayesian Gaussian mixture models

To generate training data from the simulated policies, we need to be able to sample from a probability distribution of the state space that was visited in the previous iteration. To do so, we apply Bayesian Gaussian mixture modelss. Mixture of Gaussians are a form of *unsupervised machine learning* (see, e.g., Murphy (2012)), and are commonly applied to approximate probability distributions from observed data. Suppose that we have m data samples $\mathbf{X} = \{\mathbf{x}_i : 1 \leqslant i \leqslant m\}$. One then can approximate any probability density by $\rho \approx \rho_{estimated}$ as a mixture of Gaussians:

$$\rho_{estimated}(\mathbf{x}) = \sum_{l=1}^{L} \pi_l \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_l, \mathbf{\Sigma}_l), \tag{26}$$

where the mean vectors $\boldsymbol{\mu}_l \in \mathbb{R}^N$, the covariance matrices $\boldsymbol{\Sigma}_l \in \mathbb{R}^{N \times N}$, the weights π_l with $\sum_{l=1}^{L} \pi_l = 1$, and the number of components L are fitted to $\mathbf{X}$ (see, e.g., Rasmussen (2000) and Blei and Jordan (2005)).

References

Bengio, Y., O. Delalleau, and N. Le Roux. "The curse of highly variable functions for local kernel machines". In: *Neural Information Processing Systems*. 2005.

Bewley, Truman F. *Fiscal and Monetary Policy in a General Equilibrium Model*. Cowles Foundation Discussion Papers 690. Cowles Foundation for Research in Economics, Yale University, January 1984.

Blei, David M. and Michael I. Jordan. "Variational inference for Dirichlet process mixtures". In: *Bayesian Analysis* 1 (2005), pp. 121–144.

Brumm, Johannes, Dominika Kryczka, and Felix Kubler. "Recursive Equilibria in Dynamic Economies With Stochastic Production". In: *Econometrica* 85.5 (2017), pp. 1467–1499.

Brumm, Johannes, Dmitry Mikushin, Simon Scheidegger, and Olaf Schenk. "Scalable high-dimensional dynamic stochastic economic modeling". In: *Journal of Computational Science* 11 (2015), pp. 12–25.

Brumm, Johannes and Simon Scheidegger. "Using Adaptive Sparse Grids to Solve High-Dimensional Dynamic Models". In: *Econometrica* 85.5 (2017), pp. 1575–1612.

Cho, In-Koo and Thomas Sargent. "Self-confirming equilibrium". In: *The New Palgrave Dictionary of Economics*. Ed. by Steven Durlauf and Larry Blume. 2008.

Citanna, Alessandro and Paolo Siconolfi. "Recursive equilibrium in stochastic OLG economies: Incomplete markets". In: *Journal of Mathematical Economics* 48.5 (2012), pp. 322–337.

Constantine, Paul G. *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2015.

Constantine, Paul G., Eric Dow, and Qiqi Wang. "Active Subspace Methods in Theory and Practice: Applications to Kriging Surfaces". In: *SIAM Journal on Scientific Computing* 36.4 (2014), A1500–A1524. eprint: `https://doi.org/10.1137/130916138`.

Dow, Eric and Qiqi Wang. "Output based dimensionality reduction of geometric variability in compressor blades". In: *51st AIAA Aerospace Sciences Meeting* (2013).

Feldman, Mark and Christian Gilles. "An expository note on individual risk without aggregate uncertainty". In: *Journal of Economic Theory* 35.1 (1985), pp. 26–32.

Golub, Gene H. and Charles F. Van Loan. *Matrix Computations (3rd Ed.)* Baltimore, MD, USA: Johns Hopkins University Press, 1996.

Judd, Kenneth L, Lilia Maliar, Serguei Maliar, and Rafael Valero. "Smolyak method for solving dynamic economic models: Lagrange interpolation, anisotropic grid and adaptive domain". In: *Journal of Economic Dynamics and Control* 44 (2014), pp. 92–123.

Krueger, Dirk and Felix Kubler. "Computing equilibrium in OLG models with stochastic production". In: *Journal of Economic Dynamics and Control* 28.7 (2004), pp. 1411–1436.

Kubler, Felix and Karl Schmedders. "Stationary Equilibria in Asset-Pricing Models with Incomplete Markets and Collateral". In: *Econometrica* 71.6 (2003), pp. 1767–1793.

Ljungqvist, L. and T.J. Sargent. *Recursive macroeconomic theory*. Mit Press, 2012.

Lukaczyk, Trent W, Paul Constantine, Francisco Palacios, and Juan J Alonso. "Active subspaces for shape optimization". In: *10th AIAA Multidisciplinary Design Optimization Conference*. 2014, p. 1171.

Murphy, Kevin P. *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012.

Rasmussen, Carl Edward. "The Infinite Gaussian Mixture Model". In: *In Advances in Neural Information Processing Systems 12*. MIT Press, 2000, pp. 554–560.

Rasmussen, Carl Edward and Christopher K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.

Scheidegger, Simon and Ilias Bilionis. "Machine Learning for High-Dimensional Dynamic Stochastic Economies". In: *Preprint available at SSRN: https://ssrn.com/abstract=2927400* (2017).

Figure 1: Linear forecasts

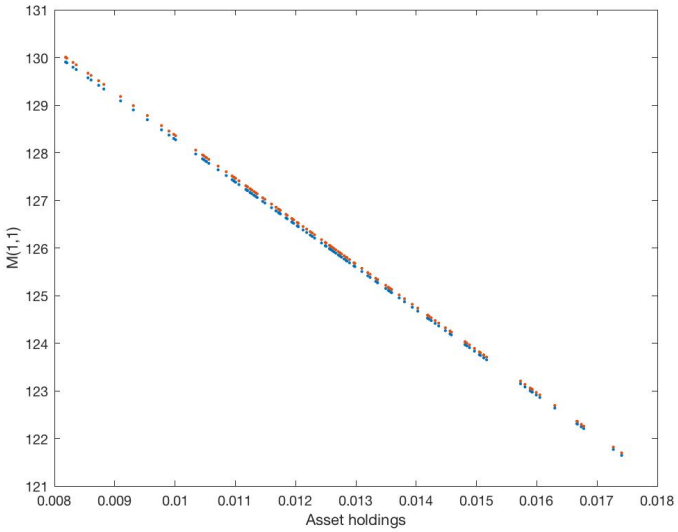

Figure 2: Linear forecasts in nonlinear world